SONDERDRUCK FÜR
sogeti
Part of Capgemini

Ausgabe 01 | 2020
Deutschland € 12,90 Österreich € 13,90 Schweiz sfr 22,20
4 194164 112900 01

OBJEKTspektrum

Software-Architektur und IT für Profis

www.OBJEKTspektrum.de

Moderne Vorgehensmodelle – Was auf Scrum und Kanban folgt

Interview mit Oliver Zeigermann

Der Architekt und Experte für Machine Learning erzählt von den Anfängen der aktuellen KI-Euphorie, differenziert die verschiedenen Varianten und räumt ein, dass ihn KI zum Denken bringt

- Erfahrungsbericht: Von der eigenen Maschine in die Cloud
- WebAsssembly ergänzt Ökosystem rund um JavaScript
- Softwarearchitektur: Den richtigen Architekturstil finden

G 6540F

Auswirkungen von KI auf den Softwaretest

Testen von und mit Künstlicher Intelligenz

Dieser Artikel beschäftigt sich mit den neuen Herausforderungen beim Testen intelligenter Maschinen, aber auch damit, wie wir mithilfe Künstlicher Intelligenz die Qualität der IT-Systeme vorhersagen und so handeln können, dass ein Rückgang der Qualität angegangen wird, bevor der Anwender Fehler irgendeiner Art wahrnimmt. Auf diese Weise stellen wir sicher, dass die Stakeholder nachhaltig der Qualität ihrer IT-Systeme vertrauen können und ihren gewünschten Mehrwert auch tatsächlich erzielen.

„Hey Siri, lies meine Textnachrichten vor", sprach ich und mein iPhone erzählte mir sofort, welche Textnachrichten ich gerade erhalten hatte. Noch vor zehn Jahren hätte es niemand geglaubt, wenn man solche Möglichkeiten vorhergesagt hätte. Heute hingegen wundert sich keiner mehr, wenn Sie sich mit einem Gerät unterhalten.

Verwendet Ihre Organisation bereits einen Chatbot? Wird er von Ihren Kunden angenommen? Und erhöht er tatsächlich den Wert Ihres Unternehmens? Und falls die Kunden ihn nicht mögen: Was hätten Sie besser machen können?

Die neuen Möglichkeiten, die intelligente Maschinen heute bieten wie Roboter, Chatbots, aber auch digitale Assistenten wie Siri, Alexa und Google Assistant, sind immens. Sie gehen jedoch auch mit neuen und unerwarteten Risiken einher.

Um den neuen Gefahren auf die Spur zu kommen, stellen wir in diesem Artikel eine Erweiterung der ISO25010-Qualitätsmerkmale [SQuaRE] vor. Diese ermitteln beispielsweise, ob die Persönlichkeit Ihres Chatbots zu Ihrer Zielgruppe passt, zum Beispiel bei so einfachen Dingen wie dem Ansprechen mit „Du" oder „Sie".

Wir erörtern auch, wie sich intelligente Maschinen für das Testen einsetzen lassen. KI-unterstützte Tools überfluten jetzt schon den Markt. Die meisten sind indes sehr einfach und definitiv noch nicht imstande, menschliche Tester vollkommen zu ersetzen. Unsere Vision ist es dennoch, dass KI-unterstütztes Instrumentarium großen Nutzen bei der Verbesserung der Qualität von IT-Systemen stiftet.

Derzeit führen wir die Tests in einer Testumgebung durch, die große Mengen an nützlichen Daten erzeugt. Beim kontinuierlichen Testen hören wir nicht auf, Daten über die Qualität des Systems zu sammeln, und zwar auch, nachdem das System in Betrieb genommen wurde. Die Überwachung der Live-Umgebung erweitert die Datenmenge über die operative Qualität des IT-Systems.

Und diese Daten bieten nach unserer Vision die wichtigsten neuen Möglichkeiten bei der Nutzung von Künstlicher Intelligenz für das Testen: Mit ihnen können wir einen Schritt in Richtung Qualitätsprognose vollziehen. Anhand der Qualitätsprognose kann das IT-Team erkennen, wo die Qualität abnehmen wird, und mit der Störungsbehebung beginnen, bevor irgendein Anwender einen Ausfall beobachtet hat. So hilft uns die Künstliche Intelligenz, den Wert der IT für den Betrieb in unseren Organisationen tatsächlich zu steigern.

„Warum brauchen wir Überwachung und Prognose? Wir haben das doch in der Vergangenheit auch nicht gebraucht", werden einige Kritiker einwenden. Doch wir wissen alle, dass die IT-Welt sehr schnell voranschreitet. Künstliche Intelligenz (KI) durchdringt in einem atemberaubenden Tempo unseren Lebensalltag. KI befähigt uns, Dinge mit unseren IT-Systemen zu tun, die zuvor undenkbar waren. Sie stellt uns andererseits aber auch vor viele neue Herausforderungen.

Den IT-Prozess im Wertebeitrag-Modell erfassen

In der IT-Welt von heute ist es letztendlich das Ziel der IT, geschäftlichen Mehrwert zu generieren. Dieser geschäftliche Mehrwert kann finanzieller Natur sein oder einer Art moralischer Natur. Um diesen Mehrwert erreichen zu können, gibt es eine Reihe von Zielen, an denen es zu arbeiten gilt (**siehe Abbildung 1**).

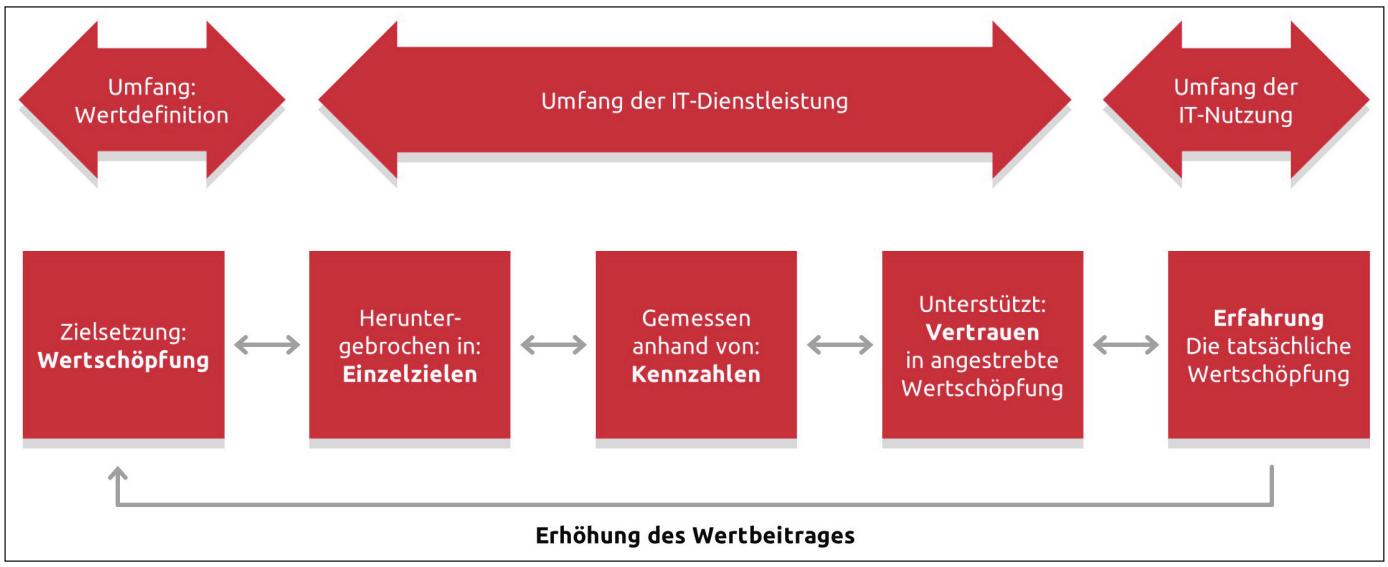

Abb. 1: Erhöhung des Wertebeitrages (Bild: Sogeti)

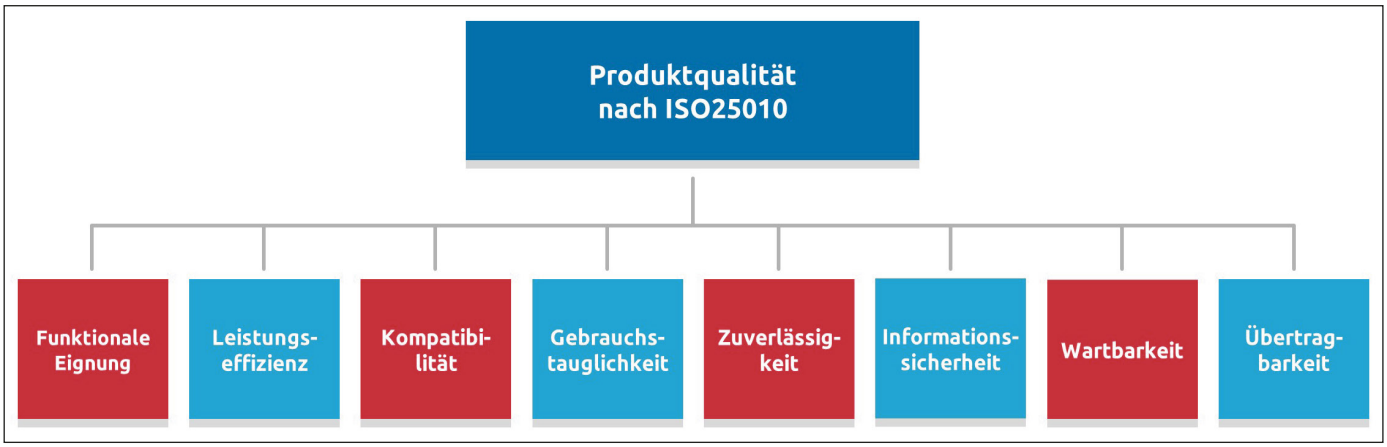

Abb. 2: Qualitätsmerkmale für Produktqualität nach ISO25010

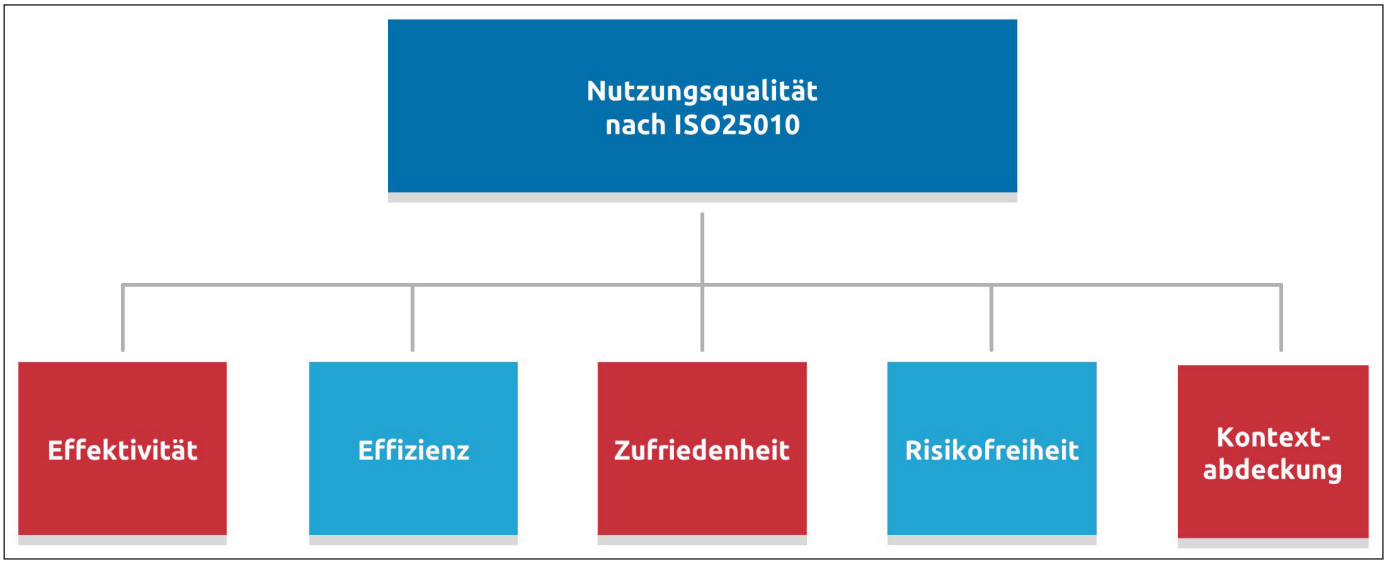

Abb. 3: Qualitätsmerkmale für Nutzungsqualität nach ISO25010

Damit wir wissen, ob wir diese Ziele erreichen, definieren wir Kennzahlen. Das Messen dieser Indikatoren erfolgt durch das Testen. Das Ergebnis des Testens sind Informationen, die dazu beitragen, dass die Akteure das Vertrauen in die eigene Fähigkeit haben, den erstrebten Mehrwert zu realisieren. Ist ihr Vertrauen groß genug, wird das System produktiv gesetzt und die Akteure erfahren den tatsächlichen Wert. Aber während dieser operativen Nutzung wollen wir weiterhin die Kennzahlen überwachen.

Im Test- und Überwachungsprozess können wir uns intelligenter Maschinen bedienen, damit sie Tätigkeiten übernehmen und den Prozess effizienter und effektiver gestalten. Wir nennen das „Testen *mit* intelligenten Maschinen". Wir müssen uns aber auch das „Testen *von* intelligenten Maschinen" anschauen. Denn das unterscheidet sich vom Testen herkömmlicher IT-Systeme, beispielsweise, weil der korrekte Output, den eine lernende Maschine heute liefert, ein anderer ist als der ebenso korrekte Output, den sie morgen liefert.

Testen VON intelligenten Maschinen: Qualitätskennzahlen auf der Grundlage von Qualitätsmerkmalen

Um ein klares Bild vom Qualitätsniveau eines jeden Systems zu erhalten und festzulegen, was in welcher Intensität getestet werden soll, müssen wir zwischen einigen Teilaspekten der Qualität unterscheiden. Dazu verwenden wir Qualitätsmerkmale. Einen weit bekannten Standard für Qualitätsmerkmale bietet die ISO-Norm 25010 für Softwarequalität mit einer Liste von 8 Hauptmerkmalen für Produktqualität und 5 Hauptmerkmalen für die Nutzungsqualität.

Mit dem schnellen Aufkommen der KI und Robotik reicht aber diese Liste heute nicht mehr aus. Um über die richtigen Testarten entscheiden zu können, die erforderlich sind, um KI korrekt zu testen, bedarf es zusätzlicher Qualitätsmerkmale. Schauen wir uns zunächst die bestehenden Qualitätsmerkmale an. Die 8 Hauptmerkmale für Produktqualität nach ISO25010 zeigt **Abbildung 2**. Die Nutzungsqualität bezieht sich auf das Ergebnis der menschlichen Interaktion mit dem IT-System. Diese Nutzungsqualität unterscheidet zwischen 5 Hauptmerkmalen (**siehe Abbildung 3**).

Zusätzliche Qualitätsmerkmale für das Testen intelligenter Maschinen

Heute sehen wir Maschinenintelligenzsysteme, die über viel mehr Optionen verfügen. So wird Input häufig anhand von Sensoren gesammelt (z. B. Geräte im Bereich Internet der Dinge) und der Output kann in physischer Form erfolgen (wie die Bewegung von Artikeln in einem Lager). Das verlangt nach einer Erweiterung der ISO25010-Liste der Software-Qualitätsmerkmale. Diese neuen Qualitätsmerkmale werden in den folgenden Abschnitten beschrieben.

Wir haben drei neue Gruppen von Qualitätsmerkmalen hinzugefügt: intelligentes Verhalten, Moral und Persönlichkeit

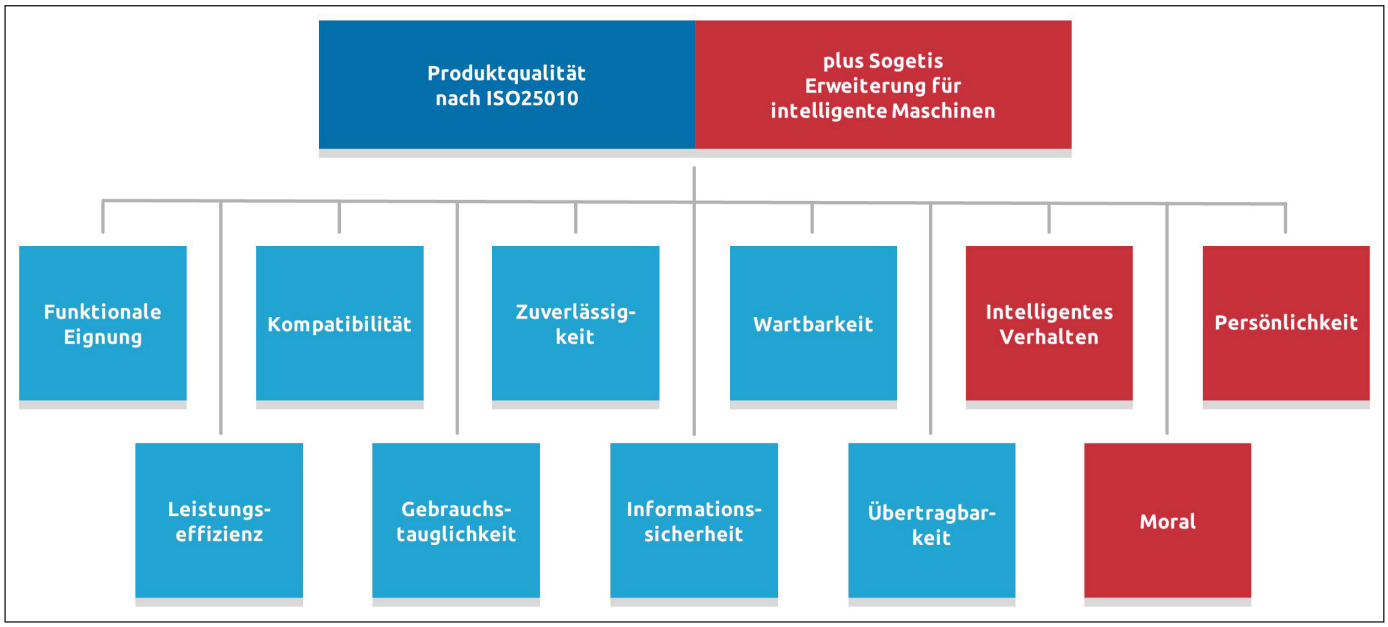

Abb. 4: Qualitätsmerkmale für Produktqualität nach ISO25010, erweitert um Sogetis Merkmale für intelligente Maschinen

(**siehe Abbildung 4**). In den jeweiligen Abschnitten beschreiben wir diese Hauptmerkmale und ihre Untermerkmale.

Intelligentes Verhalten

Intelligentes Verhalten ist die Fähigkeit, zu verstehen. Es ist im Grunde eine Kombination aus logischem Denken, Gedächtnis, Fantasie und Urteilsbildung. Jede dieser Fähigkeiten stützt sich auf die anderen. Intelligenz ist eine Kombination aus kognitiven Fähigkeiten und Wissen, die durch anpassungsfähiges Verhalten deutlich wird.

Das Qualitätsmerkmal „intelligentes Verhalten" besteht aus den folgenden Untermerkmalen:

Lernfähigkeit: Es handelt sich um die Fähigkeit, zu verstehen und Erfahrungen zu nutzen. Wie lernt eigentlich eine intelligente Maschine? Wir unterscheiden hierbei drei Lernebenen. Die erste Ebene ist das auf Regeln beruhende Lernen. Die zweite Ebene basiert auf dem Sammeln und Deuten von Daten und – davon ausgehend – Lernen über eine Umgebung. Auf der dritten Ebene erfolgt das Lernen durch das Beobachten des Verhaltens anderer und das Imitieren dieser Verhaltensweisen.

Improvisation: Passt sich die intelligente Maschine an neue Situationen an? Improvisation ist die Fähigkeit eines intelligenten Systems, in neuen Situationen die richtigen Entscheidungen zu treffen. Situationen, die keinen Bezug zu gesammelten Erfahrungen haben, erfordern eine schnelle Interpretation der neuen Informationen und die Anpassung bereits bestehender Verhaltensweisen.

Transparenz der Entscheidungsfindung: Kann ein in der Situation involvierter Mensch nachvollziehen, wie eine Maschine zu ihrer Entscheidung kommt? Es muss transparent sein, wie ein KI-System Entscheidungen trifft, welche Datenpunkte ausschlaggebend sind und wie sie gewichtet werden. Transparenz bedeutet auch Vorhersehbarkeit. Es ist wichtig, dass ein Roboter so antwortet, wie es die Menschen erwarten, die mit dem Roboter zusammenarbeiten.

Zusammenarbeit: Wie gut arbeitet der Roboter mit seinem menschlichen Kollegen zusammen? Versteht er das erwartete und unerwartete Verhalten von Menschen? Wie gut Kommunikation in einem Team aus Menschen und Robotern funktioniert, ist von entscheidender Bedeutung.

Natürliche Interaktion: Interaktion ist wichtig, sowohl bei verbaler als auch bei nonverbaler Kommunikation. Vor allem bei sozialen Robotern ist es wichtig, dass die Art und Weise, wie Menschen mit einem Roboter interagieren, natürlich ist.

Moral

Moral bezieht sich hier auf die Grundsätze bei der Unterscheidung zwischen richtigem und falschem beziehungsweise zwischen gutem und schlechtem Verhalten. Ein sehr bekannter Science-Fiction-Autor, der sich sehr intensiv mit der Moralität intelligenter Maschinen befasst hat, ist Isaac Asimov. In einem seiner Beiträge stellt er in den „Gesetzen der Robotik" die moralischen Grundsätze auf, die intelligente Maschinen befolgen sollten.

Leider müssen wir feststellen, dass – anders als in Asimovs Geschichten – diese Robotergesetze in die allermeisten intelligenten Maschinen nicht eingebaut sind. Es liegt an den Teammitgliedern in der Rolle von digitalen Testanalysten zu beurteilen, inwieweit sich die intelligente Maschine an diese Gesetze hält.

Das Qualitätsmerkmal „Moral" besteht aus den folgenden Untermerkmalen:

Ethik: Bei der Ethik geht es darum, nach gewissen Prinzipien zu handeln. Die Basis bilden Gesetze, Regeln und Vorschriften, gleichwohl ist die Beachtung von ungeschriebenen moralischen Werten bei der Ethik das Wichtigste.

Privatsphäre: Hier geht es um den Zustand von unerwünschtem oder ungebührlichem Eindringen in das Privatleben beziehungsweise der Beeinträchtigung der Privatangelegenheiten eines Menschen und dass sich die intelligente Maschine an den Schutz des Privatlebens hält.

Freundlichkeit zu Menschen: Dieses Merkmal bezieht sich darauf, dass intelligente Maschinen den Menschen oder der Menschheit im Allgemeinen keinen Schaden zufügen dürfen. Im heutigen Diskurs über die allgemeine Einstellung der KI gegenüber der menschlichen Gesellschaft wird häufig von „beneficial AI" (wohltätiger KI) gesprochen.

Persönlichkeit

Die Persönlichkeit ist die Kombination von Eigenschaften und Fähigkeiten, die ein Individuum auszeichnen. Das Qualitätsmerkmal „Persönlichkeit" besteht aus den folgenden Untermerkmalen:

Stimmung: Die Stimmung ist ein vorübergehender Geisteszustand oder Gefühl. Wird eine intelligente Maschine stets in derselben Stimmung sein? Gute Stimmung ist bei kollaborativen Robotern

beziehungsweise Cobots, die ihr Verhalten an das Verhalten der sie umgebenden Menschen anpassen, besonders wichtig.
Empathie: Dies ist die Fähigkeit, die Gefühle eines anderen zu verstehen und zu teilen. Maschinen können keine Empathie empfinden, es ist jedoch wichtig, dass sie Empathie simulieren. Sie sollten in der Lage sein, menschliche Emotionen zu erkennen und darauf zu reagieren.

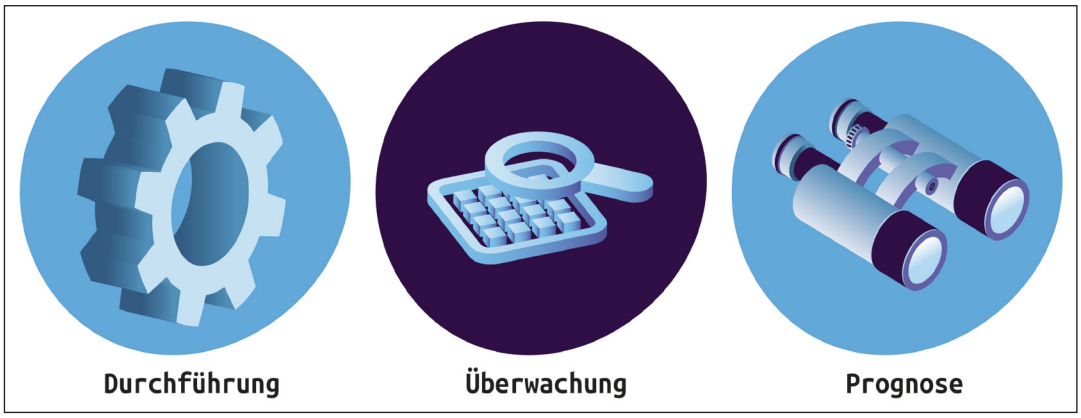

Abb. 5: Testen bewegt sich von reaktiver Aktivität (z. B. Testdurchführung) hin zur Testüberwachung (z. B. Datenüberwachung von Betriebsprodukten), eventuell wird das Testen zur Qualitätsprognose

Humor: Humor ist die Fähigkeit, der Unzulänglichkeit der Welt und der Menschen, den alltäglichen Schwierigkeiten und Missgeschicken mit heiterer Gelassenheit zu begegnen. Humor drückt sich unter anderem im geschriebenen oder gesprochenen Wort oder in der Mimik aus. Roboter verfügen über diese Fähigkeit noch nicht, aber vielleicht ändert sich das in Zukunft. Außerdem ist es eine Herausforderung, dass Roboter diese sehr spezielle Eigenart des menschlichen Verhaltens erkennen.
Charisma: Charisma ist die unwiderstehliche Attraktivität oder der Charme, der in anderen Menschen Hingabe auslöst. Ob Menschen intelligente Maschinen mögen oder sie gar lieben, liegt daran, ob sie sich als so spannend erweisen, dass Menschen gar nicht mehr von ihnen lassen wollen.

Verkörperung als Untermerkmal der Gebrauchstauglichkeit

Neben den drei zusätzlichen Hauptmerkmalen haben wir zum bestehenden Hauptmerkmal „Gebrauchstauglichkeit" der ISO25010-Norm noch ein Untermerkmal hinzugefügt. Dieses neue Untermerkmal ist die „Verkörperung". Verkörperung stellt einfach die Frage: „Sieht es angemessen aus?"
Mit physischen Robotern ist es wie mit der Benutzeroberfläche von Chatbots und sogar mit Smart Speakers (vernetzte Lautsprecher): Es ist von großer Wichtigkeit, wie sie aussehen und wie sie sich in die Umgebung fügen, in der sie arbeiten sollen. Ein wichtiger Punkt hierbei ist, dass die äußerliche Erscheinung der intelligenten Maschine zu der Funktion passen muss, welche sie erfüllen soll.

Testen MIT intelligenten Maschinen: auf dem Weg zur Qualitätsvorhersage

Im digitalen Zeitalter bilden die auf den Qualitätsmerkmalen beruhenden Kennzahlen den Rahmen für das Testen. Eine Organisation muss wachsen und sich entwickeln, um die Vermarktungszeit zu verkürzen, die wachsenden Datenmengen zu bewältigen und komplexe neue Technologien wie KI einzusetzen. Damit beginnt für eine Organisation die Reise hin zur Qualitätsvorhersage.

Qualitätssicherung und Testen durchlaufen einen Wandel von einer reaktiven Tätigkeit, oft mit einem Schwerpunkt auf Testdurchführung, hin zu einer aktiven Überwachung von Daten aus operativen Systemen (siehe **Abbildung 5** und [Ven18]). Die Daten, die über die Qualität eines IT-Systems gesammelt werden, lassen sich als Input für KI-unterstützte Modelle verwenden, die schlussendlich die Qualität prognostizieren werden. Auf diese Weise können Fehler vorhergesagt und korrigiert werden, noch bevor irgendein Nutzer auf einen Fehler stößt.
Durch die steigende Nachfrage nach schneller Softwarelieferung werden Softwarelebenszyklusmodelle wie Agile und DevOps heute schon allgemein verwendet. Aber wie lässt sich diese Geschwindigkeit noch weiter erhöhen?
Die Evolution des Testens lässt sich in vier Phasen unterteilen (**siehe Abbildung 6**): vom klassischen Testen, wo die Zyklen in Monaten gemessen wurden, über das agile Testen mit wochenlangen Testzyklen, zum kontinuierlichen Testen mit Zykluszeiten von mehreren Tagen und schließ-

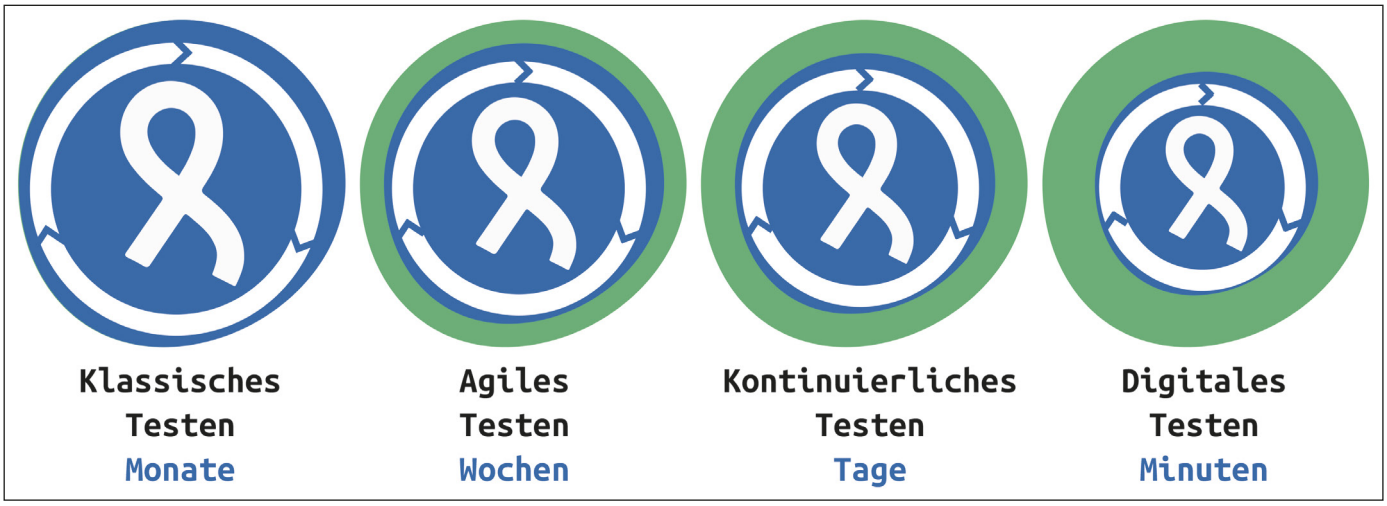

Abb. 6: Weiterentwicklung vom klassischen Testen zur Continuous Integration und Testen innerhalb von Minuten

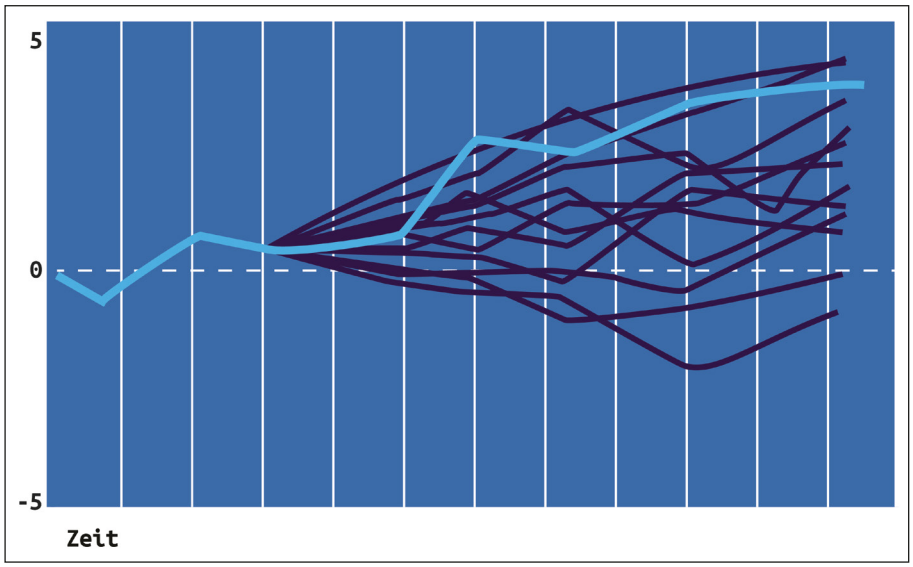

Abb. 7: Die sogenannte „Feder"-Prognose, aus der die Unsicherheit von Zeitaufwand für Mehrmodell-Ergebnisse ersichtlich ist

lich zum digitalen Testen mit Zykluszeiten im Minutenbereich.
Schlussendlich übernehmen selbstlernende und selbsterforschende Testwerkzeuge den größten Teil der Denkarbeit. Technologien wie KI können dem Menschen beim Testen moderner und komplexer IT-Lösungen helfen. Insbesondere kann ein evolutionärer Algorithmus verwendet werden, um Testfälle zu erzeugen. Zusammen mit der Nutzung von Mehrmodellprognosen für Testumgebungen verwandeln sich Testfälle in smarte Testfälle. Das Testen bleibt jedoch eine reaktive Tätigkeit. Wir müssen hingegen zu einer proaktiven Vorhersage kommen.
Qualitätsvorhersage zielt darauf ab, den Testergebnissen und Real-Life-Ergebnissen voraus zu sein. Noch bevor eine unerwünschte Situation eintritt, identifizieren die digitale Testanalyse und Überwachung bereits die sich anbahnende Fehlerwirkung, und das Team kann den zugrunde liegenden Mangel beseitigen. Um in diese Lage zu kommen, müssen zahlreiche Voraussetzungen erfüllt sein. Die Tests müssen mit automatisierten Werkzeugen generiert werden, beispielsweise mit modellbasiertem Testen und digitalen Zwillingen. Die Testdurchführung muss größtenteils automatisiert erfolgen, um die rechtzeitige Sammlung von Daten aus den Testergebnissen zu gewährleisten. Und die Daten aus der Testdurchführung und der Überwachung müssen automatisch gesammelt und vorverarbeitet werden.

Wie funktioniert Qualitätsvorhersage in der Praxis?

Der Prozess der Qualitätsvorhersage (engl. Quality Forecasting) beginnt mit den Daten, die während der Testdurchführung und der Live-Überwachung gesammelt werden. Diese Daten werden sodann von KI-unterstützten Modellen interpretiert. Ähnlich wie bei der Wettervorhersage verlässt man sich nicht nur auf ein einziges Modell. Es sind vielmehr zahlreiche Modelle erforderlich, die jeweils unterschiedliche Parameter nutzen, um zu prognostizieren, wie ein IT-System funktionieren wird.
Die Modelle haben jeweils einen unterschiedlichen Grad an Unsicherheit in Bezug auf das zukünftige Verhalten eines Systems von Systemen, so wie bei der Wettervorhersage die Wetterentwicklung im Zeitverlauf unterschiedliche Ergebnisse für jedes Modell anzeigt. Im Laufe der Zeit wird die gezeigte Bandbreite der Ergebnisse immer größer (**siehe Abbildung 7**).
Dieses Konzept lässt sich sehr gut auf komplexe IT-Landschaften übertragen. Wir entwickeln ständig neue Umgebungsmodelle und verwenden dabei einen evolutionären Algorithmus, um extreme Testfälle zu erstellen, die belegen, dass die Nutzung einer IT-Lösung immer noch sicher ist, das heißt, es werden keine kritischen Grenzen des Systems überschritten. Man kann sogar so weit gehen und eine Reihe von Vorhersagemodellsituationen erzeugen, in denen wir versuchen, diese Extreme zu verursachen.
Aber wie weiß man, wie die Prognose tatsächlich aussehen sollte? Wie bei der kurzfristigen Wettervorhersage sind das Wissen und die Intuition des Menschen weiterhin vonnöten, um die Ergebnisse der Mehrmodellprognosen zu interpretieren. Der Testanalyst entwickelt sich zum Qualitätsprognostiker der zukünftigen IT-Systeme!

Fazit

Die Einführung von KI, maschinellem Lernen und anderer Formen intelligenter Maschinen in unseren Geschäftsprozessen stellt uns vor neue Herausforderungen bei der Qualitätssicherung und dem Testen. Das Messen der Kennzahlen, das nötig ist, um diese neuen Herausforderungen zu meistern, erfordert eine Erweiterung der ISO25010-Qualitätsmerkmale um neue Attribute für intelligentes Verhalten, Moral und Persönlichkeit.
Und die KI verleiht uns neue Fähigkeiten, zum Beispiel die Qualität von IT-Systemen zu prognostizieren und so zu handeln, dass ein Rückgang der Qualität angegangen wird, noch bevor der Nutzer Fehlerwirkungen irgendwelcher Art feststellt. So gewährleisten intelligente Maschinen, dass die Akteure der Qualität ihrer IT-Systeme nachhaltig vertrauen können und den angestrebten Mehrwert tatsächlich realisieren. ||

Literatur & Links

[SQuaRE] ISO/IEC 25010:2011, System and Software-Engineering – Qualitätskriterien und Bewertung von System und Softwareprodukten (SQuaRE) – Qualitätsmodell und Leitlinien, International Standards Organization

[Ven18] T. van de Ven, R. Marselis, H. Shaukat, Testing in the digital age. AI makes the difference, SogetiBooks, 2018

Der Autor

Rik Marselis
(rik.marselis@sogeti.com)
arbeitet als Testexperte bei Sogeti in den Niederlanden. Er ist Mitglied von Sogeti-Labs, dem internationalen F&E-Netzwerk. Rik Marselis ist zudem Autor zahlreicher Buchveröffentlichungen.